(2)

# AIR WAR COLLEGE

# RESEARCH REPORT

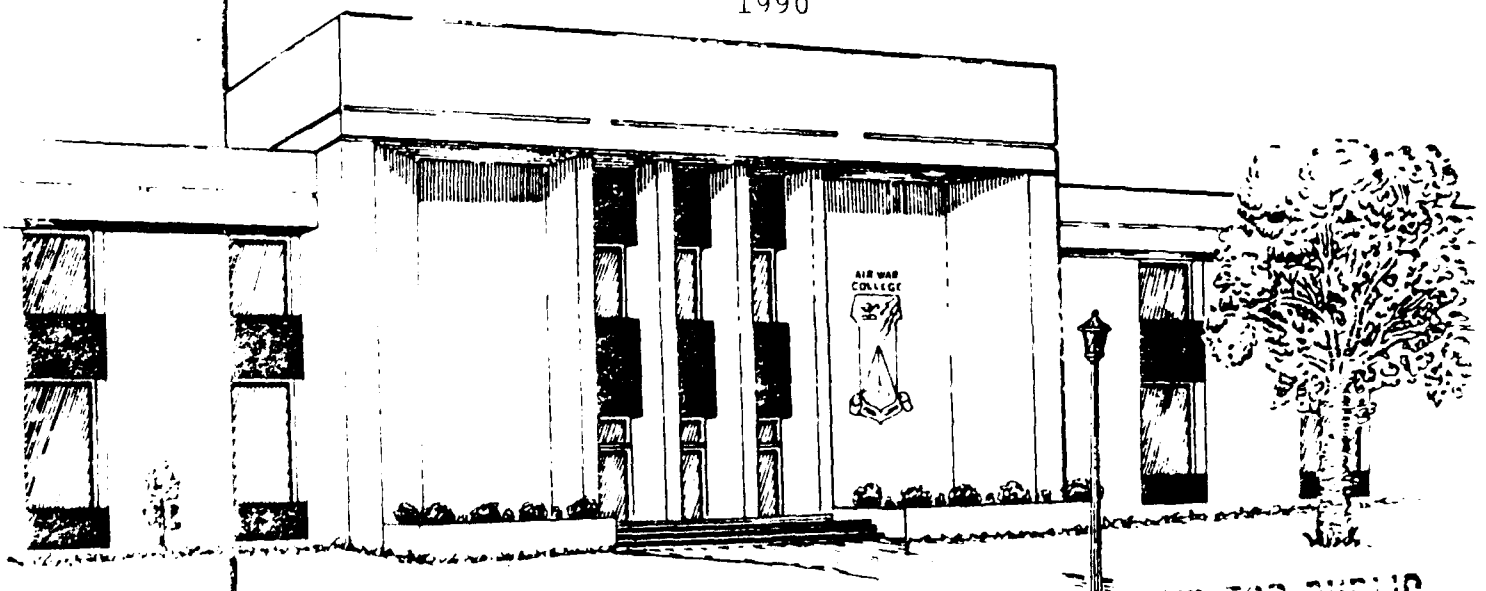


DATA AND INFORMATION INTEGRITY IN A DISTRIBUTED ENVIRONMENT



MR D. CARL ABERNETHY, JR.

1990



AIR UNIVERSITY
UNITED STATES AIR FORCE
MAXWELL AIR FORCE BASE, ALABAMA

AIR WAR COLLEGE

AIR UNIVERSITY

# DATA AND INFORMATION INTEGRITY IN A DISTRIBUTED ENVIRONMENT

by

D. Carl Abernethy, Jr.
Civilian

A DEFENSE ANALYTICAL STUDY SUBMITTED TO THE FACULTY

IN

FULFILLMENT OF THE CURRICULUM

REQUIREMENT

Advisor: Bruce T. Morland, Jr.

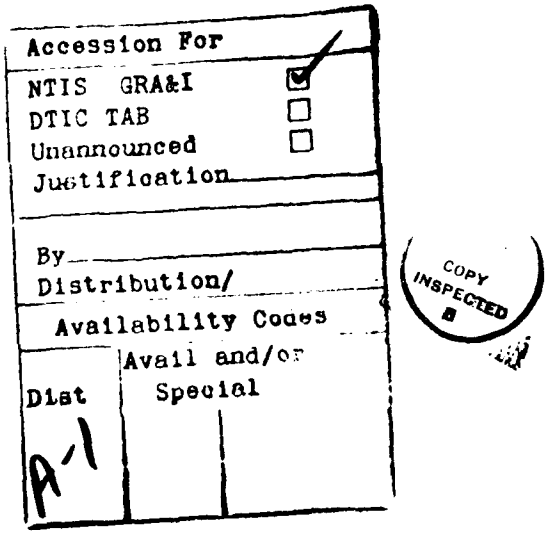
MAXWELL AIR FORCE BASE, ALABAMA

MAY 1990

## DISCLAIMER

This study represents the views of the author and does not necessarily reflect the official opinion of the Air War College or the Department of the Air Force. In accordance with Air Force Regulation 110-8, it is not copyrighted but is the property of the United States government.

Loan copies of this document may be obtained through the interlibrary loan desk of Air University Library, Maxwell Air Force Base, Alabama 36112-5564 (telephone (205) 293-7223 or AUTOVON 875-7223).



## EXECUTIVE SUMMARY

TITLE: Data and Information Integrity in a Distributed Environment.

AUTHOR: D. Carl Abernethy, Jr.

The 90's will be the era of Information Management in Computer Processing. Information Management demands the integrity of data and information that we process and handle. As we move into this new age, we are losing the ability to ensure integrity in a distributed processing environment. This is due in part to the proliferation of terminals, workstations and the advent of networking as we move from a centralized approach to data processing and data-basing.

Integrity is more than a security issue. It encompasses accuracy, correctness, and validity of data. Database development, database management systems, networking of terminals and systems, and the distributed environment of software and information compound integrity concerns.

Until we recognize that information is our most precious resource, we will ignore the importance of integrity concerns and their impact on the computer world. This paper will address these issues.

## BIOGRAPHICAL SKETCH

Mr. Abernethy has been an employee of the Department of Defense, NSA, for 26 years, and holds the professional certification as a Computer Systems Analyst. Within the computer science career field, his experience and expertise centers upon software development, specifically, in the area of applications development vice systems software.

In his most recent assignment, Mr. Abernethy co-chaired a Data Planning Element that was responsible for developing a focus plan for ADP development. He co-authored the publication "Data Distribution Strategies for the 1990's." While, the planning element produced reports that ultimately resulted in the reorganization of ADP support for the largest analytical group at the Agency.

Mr. Abernethy's background includes working as a programmer developing analytical applications and as a computer systems analyst designing systems and databases in support of the Agency's mission. He served as a senior computer scientist in several analytical groups, responsible for operational development and support of applications and systems for those groups. He also served the Agency as a Division manager for the Computer Resources Management and Programming Division, an Analytic Support Division, and a Special Processing Support Division. He has been recognized for his accomplishments in each of these areas. He is currently a student at the Air War College.

TABLE OF CONTENTS

# CHAPTER I

## INTRODUCTION

This is the information age. With the advent of computers, computer systems, the proliferation of terminals and workstations, networking and distributed processing, the need to ensure the integrity of our data and information is becoming critical. We are being inundated with millions of bits of data that computers can store and pass at ever increasing speeds. To alter the arrangement of or remove any of these pieces of data destroys the original thought, conveying a different meaning or none at all. Perhaps the easiest way to remind ourselves of the importance of integrity is to reflect on the childhood game of "telephone." A number of people stand in a circle. The first person whispers a story to the person to his right and this process is repeated until the story is passed around the circle and repeated back to the originator. We all know that the story is changed or may be completely different. If the meaning and literal constructs of the story could be maintained without change as it passes from one person to the next regardless of the number of persons in the chain, then we have information integrity.

Although computer security has been an important requirement in the military since computer use began, it has been

only explicitly recognized in nonmilitary government and business since the late 1960's. Data and information integrity, on the other hand, are just now being recognized as part of the security environment. Additionally, the disclosure of information to someone not authorized to see it is a major focus of governmental and military security and has been a concern since long before the invention of computers. The "Department of Defense Trusted Computer System Evaluation Criteria" documents the requirements for secure computer systems at various levels. Still, no standards or guidelines are forthcoming in the area of integrity, even though it is recognized as a problem.

To protect our most precious resource adequately, we need to understand integrity and its implications and constraints in several environments. This paper will provide a working knowledge of integrity, its relationship to security, databases and database management systems, the impact of integrity in networks and the distributed environment.

## CHAPTER II

## INTEGRITY

What is data integrity and information integrity? To grasp these ideas and the seriousness of the problems we face, one must understand what we mean by integrity in both contexts. Most people in the discussion of integrity interchange the words data and information. We should understand the distinctions between these words. Yet, in the discussion of integrity to follow, these distinctions are not critical to understand the importance of data and information integrity. I also will take the liberty of interchange once the distinctions are clear. To continue to differentiate between these terms will be of little benefit to the reader as we develop an understanding of integrity in the computer environment.

For my purposes, I will consider a datum as the lowest or smallest element of an information chain. By that I mean, a datum will represent the smallest physical element of unique value within a frame of reference. For example, if we consider the written words of this article, the information chain would be the article, a chapter, a paragraph, a sentence, and a word. Words then become the data elements within this frame of reference. By chaining words together we create information to convey different thoughts. Yet,

in a different context, we know that letters comprise words and within a computer the binary representation of a letter consists of a unique string of 1's and 0's. In a graphical sense, hundreds of dots which when physically aligned in an agreed upon format represent a pictorial pattern we call a 1 or 0.

What then is information? Captain Jackson, Chief of the Technology Integration Office, Air War College, Maxwell AFB, Alabama, has postulated that "a common definition of information is the meaning humans assign to data. That is datum has no meaning by itself but when put together with other datum then there is meaning to the perceiver." (1) The point being, a frame of reference or agreed upon standard and level of precision must be defined and approved. Approval needs to be by all parties concerned with the data elements or groupings of those elements that comprise information.

Integrity can be defined as "those qualities which give data and systems both internal consistency and a good correspondence to real-world expectations for the systems and data. Primarily, the expectation of integrity means that systems and data remain predictably constant and change only in highly controlled and structured ways. This concept of integrity is tied to both an internal and an external consistent standard." (2:16) Another definition for data integrity is "the state that exists when computerized data is the same as that in the source documents or has been correctly computed from source data and has not been exposed to accidental or malicious alteration or destruction." (3:32)

Ronda Henning and Swen Walker expressed the best encapsulation of the integrity concept in their article "Data Integrity vs. Data Security: A Workable Compromise." They identified six functional areas:

a. How correct we think the information is,
b. How confident we are that the information is from its original source,
c. How correct the functioning of the process is,
d. How closely the process function corresponds to its designed intent,
e. How confident we are that the information in an object is unaltered, or was correctly modified, and
f. How correct the information in an object is. (4:335)

We have defined data and information. What then is the process that Henning and Walker allude to? Process can be defined as the desired intent of achieving the same results if the same code (software) executes repeatedly with the same input. We rely and expect the operation of our computer systems to run properly with known and reliable results. That the correctness of the result can be guaranteed. We classify this as system integrity. Yet, two areas of concern are immediately identifiable, the electromechanical device (the computer - hardware) and the set of instructions (code - software) that execute on the system. Manufacturing integrity and the consumer market assure us of hardware reliability. The user or consumer must then ensure that the operating environment falls within the manufactures' specifications to maintain that integrity. We now begin to understand the high cost of software development in terms of time and money. The creation of instruction sets that will produce desired, accurate results on data is demanding.

meticulous work. Slipshod procedures and lack of standards or controls induces errors in the software development process that significantly impact the integrity of the resulting product of the instruction set executed. Developers of code must be aware of the traps awaiting them and resist the temptations to produce code that does not ensure integrity of results. Again, the consumer market controls commercial software and the demands of the user control the evolution of internally developed code.

Integrity in our context embodies accuracy, correctness, and validity of data. The dominant problem of integrity is the problem of ensuring that the data is accurate. We must protect information from errors in data entry, by mistakes made by people manipulating the data or by people operating the system. Programmers make errors, systems fail, and even deliberate actions are taken to falsify data. We must maintain internal consistency. The simplest way is to prevent data modification. Given that change must occur then "the primary assurance of integrity is the knowledge of authorship." (2:17) Two other internal consistency controls are applicable. One is to constrain change through the execution of specified software that certifies the change of data only in a specific way. Secondly, to ensure change only occurs when performed by two different people authenticated to perform the change. We define this as partition of change. (2:17) Most importantly, when we consider integrity, we also must maintain external consistency. We must ensure the resultant outputs of our

processes match with the expectations and relationships of the outside world and reflect exactly those that existed outside the computer.

# CHAPTER III

## INTEGRITY CONSIDERATIONS

One basic responsibility of an organization's management is to take appropriate and reasonable measures to protect all its possessions. That responsibility must include its information assets as well. We only now are beginning to realize that it is as important to protect an item of information as it is to protect money or property.

There are typically five areas of concern that historically we consider as security related but are more specifically concerns of integrity: fraud, loss of confidentiality, inadvertent damage to data, malicious tampering, and physical damage to hardware. These apply to any type of computerized operation where integrity is important. Surprisingly enough, employees cause most of the damage in these five categories not outsiders. Yet, the press often emphasizes those few crimes perpetrated by criminals and hackers. The larger the organization, the more chance exists for problems in the area of inadvertent damage to data, malicious tampering, and possibly physical damage to the hardware. (5:60)

One example typifies a problem that led to some earlier prosecuted cases of computer fraud. The problem of automatic posting of interest rates to bank accounts without established

standards or levels of precision. Consider what happens when you divide 22 by 7. This may be represented as 3.14, 3.143, 3.1429, and it goes on. The differences may be subtle, yet, exploitation of these differences represented hundreds of thousands of dollars to clever crooks before the arrival of standards.

There has also been some concern expressed that large corporations and even the people within them will not always act with the public interest in mind.

> People who have had technical educatic _ aren't often well-versed in the ethical and social implication of how they use the technology. . . . Others take a more optimistic view. They stress the two points that it is not information technology that creates the problem but the choices that are made on how to use it, and that most problems arise because the new techniques have arisen in a framework of old institutions and attitudes. (6:21)

What position you happen to take really makes no difference. Information is important and it is as equally important that we protect it. The idea of treating information as an asset or more specifically a commodity to be valued is new. Unfortunately, we have little experience in dealing with this idea. Part of the problem stems from information having some unique characteristics.

- It can be reproduced, quickly and at low cost.
- When information is stolen, you are not usually deprived of its use. What you lose instead is the exclusive right to use confidential information.
- Information can be transported instantly to nearly anywhere.
- Its value is determined by its useful life, sometimes very brief.
- Its value does not add up. Two copies of the same information are not normally worth much more than one copy.

These unique characteristics have created many problems in a legal, social, and business system that is not yet truly geared to cope with the new order. Our institutions still are

> oriented primarily toward the commercial exchange of tangible products and services, not to the use--and misuse--of information. (6:21)

The draft paper on "Trusted DBMS Interpretation" states that "integrity quite often impacts security and that security is necessary to provide some aspects of integrity." (7:2) The discussion of integrity intuitively leads to a discussion of security. Especially, those aspects of security that involve integrity controls. In the next chapter we will consider computer security and the implications of information integrity.

# CHAPTER IV

## SECURITY CONSIDERATIONS

Computers have become indispensable to almost every form of modern business and government. This has led not only to an increase in the potential for misuse of hardware and software but computer data. "As the importance of computerized data increases for virtually every business, so does the danger to the security of that data. Data is under assault on a number of fronts, and figuring out how to protect it is getting harder and harder." (8:136) The people who create and work with computer products have the capability to alter or delete information stored in computers or to create totally new information. The security of this information, and other data stored in computers, is vital. Computer security encompasses the integrity, preservation, authorized use, and confidentiality of data. This starts with its generation, through its entry into computers, automatic and manual processing, output, storage, and finally its use. A primary motive for computer security is protection from intentionally caused loss. However, the news media frequently distorts computer crime and is quick to publicize its occurrence.

> "To a good approximation, every computer in the world is connected to every computer -- with few exceptions," says Robert Morris, chief scientist at the National Security Agency. That level of sharing brings with it both great benefits and

serious problems. Computer users can share information, resources and processing power. However, using the same links, they can destroy or alter a rival's data, eavesdrop on private communications or pass on insidious computer programs capable of proliferating like viruses, overwhelming networks and taking over computer operations. (9:199)

Now is the time to realize that we must devise a basic program that will guard our information. Computer security plays an integral role in establishing that protection. With the proliferation of personal computers (PCs) the security problem is becoming intolerable. It is common for PC users to pass around copies of software or down load programs from electronic bulletin boards. The ease with which this transfer can occur portends disaster from the insidious deployment of computer viruses. Having many PCs networked together raises the spectre of maliciousness to epidemic proportions.

Tom Manuel has identified two distinct kinds of security threat. "Besides the older, ever-present threat of equipment of software failure (and the related threat of damage done by inexperienced users), there is now a very real threat from malicious users. That problem, in turn, breaks down into two separate problems -- keeping malicious outsiders off the system, and preventing disgruntled or criminally inclined employees from attacking it." (8:137)

What aspects of security need to be addressed? The integrity issues that are directly interrelated to security will be identified. Access is the first security problem that the average user encounters. It is as necessary to limit who may use a

particular computer as it is to control the storage and retrieval of that information.

Control of the physical environment is the initial security access mechanism to the computer and ultimately the data. The next level usually involves the use of passwords. While the use of passwords has historically been thought of as an adequate security measure, traditional password systems alone no longer provide the necessary security for many commercial, government, and military activities. We now need more reliable methods to identify a specific user of a system not just someone who has access to it.

We find that many operating systems do not store passwords in encrypted files or databases. Gaining access to a system may gain you access to passwords that in turn provide access to sensitive data. This scenario may permit jumping from one system to another within a network. Access may be obtained to systems for which you have no access rights (e.g., an uncleared user may gain access to a system containing classified information or access to data of a higher classification than authorized).

Retaining a single user password for long periods of time, invites either misuse or attack. The solution would be to change passwords dynamically after each use. There is almost no opportunity to gain access and use the password later.

We also must consider the human aspect of password use. How many managers let their secretaries use their passwords to retrieve electronic mail? Do we share passwords to ease continued

operation at vacation time? How many of us routinely employ familiar personal details, such as birthdays, names of children, or simplified patterns for our passwords? These simplifications subvert the intent of password systems and make it significantly easier to access a system maliciously. It seems that the more difficult a password is to guess the more difficult it is to remember. If it is difficult to remember, we tend to write it down and again subvert the intention of password use on a computer system. Over reliance on a single system for access is foolhardy. We now have other methods in addition to password use to verify access. These are commonly called authentication systems or tests.

> If users were required to pass a combination of authentication tests, unauthorized commandeering of other users' privileges would be considerably restricted. The different methods for achieving this combination, technically known as "extended user authentication," fall into five categories. These are: something users know (like passwords), something users have (like magnetic cards - tokens), something users are (like fingerprints), something users can do (like sign their name), and someplace users are (implemented via terminal identification codes and other more secure mechanisms). (10:18)

If we are to consider the acquisition and employment of authentication technologies and systems we must be aware of the cost. The actual purchase and installation costs pale by comparison to the costs associated with defining user privileges, educating users, the life-cycle costs of maintaining the data base, and handling of problems. Given that no security system is unbeatable, authentication schemes offer a positive step forward to the traditional password systems.

Modification of information, and whether the modification results in information that is in some sense consistent or correct, are aspects of integrity. Permission to change or authorization to modify is an aspect of security control that may lead to a breach of data or information integrity. We divide authorization into two categories. The first is mandatory integrity authorization, which deals with integrity classifications reflecting importance of data, and clearances reflecting user trustworthiness. The second is discretionary integrity authorization, which we base on a user's need to modify information. Both mandatory and discretionary integrity controls can protect data from malicious tampering and destruction. These controls also protect from accidental modification and destruction through operator error or faulty software. (3:264)

Whether we consider a single workstation connected to a computer system or many workstations connected to a network, the access problems are similar. Security problems on the other hand are not. Networks offer many avenues of access to data to many people. Some of which may be sophisticated enough to subvert the security systems. Security polices must be in place to prevent the natural disaster or the malicious attack that either brings a network down or restricts access to the databases. Should each node of the network maintain its duplicate database? Distributed databasing techniques may offer a solution. Anyhow, a good security plan will include provision for disaster recovery.

In general, security systems will control, through use of specific security features, access to information. These systems ensure that only properly authorized individuals will have access to read, write, create, or delete information.

Systems that employ sufficient hardware and software integrity measures and permit processing of a range of sensitive or classified information are trusted computer systems. The Department of Defense Trusted Computer System Evaluation Criteria identifies six fundamental computer Security Requirements:

> SECURITY POLICY - There must be an explicit and well-defined security policy enforced by the system. . . . there must be a set of rules that are used by the system to determine whether a given subject can be permitted to gain access to a specific object.
> MARKING - Access control labels must be associated with objects. . . . it must be possible to mark every object with a label that reliably identifies the object's sensitivity level and/or the modes of access accorded those subjects who may potentially access the object.
> IDENTIFICATION - Individual subjects must be identified. Each access to information must be mediated based on who is accessing the information and what classes of information they are authorized to deal with.
> ACCOUNTABILITY - Audit information must be selectively kept and protected so that actions affecting security can be traced to the responsible party. A trusted system must be able to record the occurrences of security-relevant events in an audit log.
> ASSURANCE - The computer system must contain hardware/software mechanisms that can be independently evaluated to provide sufficient assurance that the system enforces the requirements previously mentioned.
> CONTINUOUS PROTECTION - The trusted mechanisms that enforce these basic requirements must be continually protected against tampering and/or unauthorized changes. (7:3)

The guidelines provided are meaningless if no one is willing to take action to implement security practices. Over the years, public ignorance of information processing and the technology

associated with the computer revolution has been an important part of security defenses. With the ever broadening base of a computer literate workforce and sophisticated users, the threat of abuse is increasing. Implementation of effective security procedures and controls will only occur with management's commitment and support. However, the most rigorous security capabilities can be undermined, even if controls and procedures do exist, but lack substance. Protection of corporate information is the responsibility of all members of the organization, and becomes more critical as the era of distributed processing is upon us. It is our responsibility to be more diligent as we build our databases and employ database management systems that control the accessibility of our information.

# CHAPTER V

## DATABASES AND DATABASE MANAGEMENT SYSTEMS

What do we mean by database or database management systems (DBMSs)? A database is a collection of information that is related or logically connected. It is organized in such a manner that data may be retrieved at will. Databases stored on computer systems are often referred to as files.

DBMSs are defined as software packages that permit multiple files to be accessed or used simultaneously. Most DBMS packages include a programming language to design specific user applications. Most will have a menu interface that allows simple constructs or databases to be created. DBMSs vary in power and depending upon their design (fixed or variable length systems) will determine many of their capabilities. Systems based on a fixed length construct will waste storage space and impose constraints on record or field size. Response time (time needed to access data) for very large databases could be slow. However, their advantage lies in ease of use and are typified by most commercial systems. Variable length systems overcome space and length limitations. They are usually more complex, requiring implementation by computer professionals or highly trained users.

In the early days of data processing, almost all data was on removable magnetic tapes or card decks. Security problems associated with unauthorized access, manipulation, or destruction were primarily of a physical nature. However, with the arrival of fixed and removable disk systems, random access processing, and remote computer access (via communications), security considerations began to increase.

Data integrity in the database management sense can be thought of as the correctness of the data itself. Also included are any associated data structures and information required to access the database. Locking mechanisms for the update and addition of information to a database are principal concerns of database integrity. If a user is updating the database, an exclusive lock mechanism must deny other users access. Specifically, if they are attempting to update the database or retrieve information.

Today's database management systems (DBMS) are essentially multilevel and multiuser data storage devices. These systems have all the potential weaknesses one might expect from a system designed for extensive user sharing. They lack a primary emphasis on security or consideration for information integrity. "The distinction between the security responsibilities of the database management system and the operating system is not well defined. The responsibilities of the database management system depend upon the security policy or the operating system." (3:249)

The installation of special safeguards provides sufficient multilevel access controls for most DBMSs but not the integrity of the data. Within the general purpose operating system environment, there are two basic types of security policies enforced. Those that provide some degree of discretionary access control, and those that provide mandatory access controls.

Discretionary access control is "a means of restricting access to objects based on the identity of subjects and/or groups to which they belong." (7:112) Systems that rely only upon discretionary security policies to provide a secure environment can be easily circumvented. For example, a user may be able to bypass the DBMS's security controls and access any database directly from the operating system. Thus permitting the database files to be read with conventional file access techniques.

Mandatory access control is "a means of restricting access to objects based on the sensitivity (as represented by a label) of the information contained in the objects and the formal authorization (i.e., clearance) of subjects to access information of such sensitivity." (7:114) Attaching labels to data and requiring clearance authorization provides protection for each user. For example, a user cannot exist at the top secret level and modify an unclassified file even if there is discretionary access to the file. Operating systems that enforce mandatory access control policies afford the DBMS all the advantages of discretionary access control, and add further security controls.

Most DBMS run on top of the computer's operating system. They normally allow the operating system to control the input and output functions for all data transfers to the storage medium. The user receives access to the databases that reside on the system for which the password is valid, only after authentication of the password.

Even where the DBMS itself provides control over access, the end use of the data cannot be controlled. Thus, most DBMSs do not provide sufficient multilevel access control. The crux of the problem is that most operating system access control mechanisms only guard the system, not the data itself. There needs to be some degree of access control at the file, record, field, and data element level for read, write, and execute permission.

> Database management systems, which first became widely available some 20 years ago, are, for many users, the single most crucial piece of software they will ever own.
>
> Minicomputer and mainframe based DBMS packages are often criticized for their incompatible data structures and inflexible user interfaces; problems commonly associated with micro-based DBMSs range from slow performance to their inability to manage sophisticated programming tasks. In addition, several areas of concern are common to both classes of DBMSs: a troubling absence of data integrity and security functions, the lack of standards, and vendors that advertise inherently nonrelational systems as relational. (11:67)

Distributed database management systems are relatively new technology and pose new integrity and security problems. We must consider the highly complicated integrity issues as an integral part of database security, operating system security, and now network security. A further discussion of databases and databasing will be addressed in the chapters concerning Networking and The Distributed Environment.

# CHAPTER VII

## NETWORKING

Computers were once scarce enough, and limited enough, that communication between them was impractical and unnecessary. Still, as the number of machines and users increased, the use and demand for communications increased. The development of high speed data communications for both local and wide area networks resulted from the ingenious use of hardware and software.

Computers have changed dramatically. Today's computers bear little resemblance to the minicomputers of the 1970's. Those machines required optimization in areas of memory management, disk input/output (I/O) control, and terminal I/O. Today's machines require optimization in network access and use.

Technology plays an important role in the security of communications systems. Local area networks (LANs), already have wide use in the United States and abroad for linking computer based systems. They represent another area of information technology in which security and integrity issues receive little attention. LANs are becoming the standard means of implementing distributed information processing systems. Securing these systems is more than a matter of convenience. It is a matter of survival. The integration of computers and teleprocessing networks has increased the scope of the problem.

When one person is using a single computer from a locked room, there is no security problem with the possible exception of electronic emanations (Tempest concerns). This observation is not useful by itself, but focuses attention on the alternate case: information security problems arise as access to computers increases. Because a network's purpose is to extend access, they inherently increase the risk to information security. It is ironic that extended access, the fundamental benefit of networking, is also the source of risk for accessing data.

It is imperative that an organization focus special attention on the network environment and carefully evaluate the risks that are unique to that environment. After identifying the risks and performing a quantitative evaluation or the vulnerabilities to loss, a cost benefit analysis will decide the protective measures which should be implemented. Preferably, this evaluation needs to be made prior to the installation of a network. Retrofitting security and integrity controls to an existing network can be expensive.

> Although network data security has improved, the security problem has not disappeared, and it's one that few network users and administrators can ignore. Security issues regarding personal computer based local area networks have become critical as more and more large corporations choose LANs as alternatives to the costly mainframes and minicomputers they once favored. (12:136)

Ten years ago the probability of a network break-in was small enough that most organizations chose to assume the risk of a break-in. They chose this alternative over an investment in security. Now, that risk has become too great to assume a passive

security posture. The frequency with which hackers breach the security of major commercial, government, and other data centers illustrates the escalation of the threat.

"While the level of security you need always depends on the application, you can analyze your network's security needs by performing a risk assessment based on the DoD's own guidelines." (12:139) Data sensitivity, specifically in the government arena, is a most important issue. "You will considerably reduce the opportunity for someone to intercept data or intrude on the system, if you can avoid any kind of remote or real-time processing on your network. . . . In the event of a breach . . . if your data is altered or obliterated or your hardware damaged, you can restore the system painlessly to current or near-current status from your backup system." (12:141)

> Virtually all of the people involved in a network are basically well-meaning and careful. The challenge is protecting them and the system from the tiny number who are malicious or foolish. Making it impossible for the latter to carry out their nefarious activities might seriously inconvenience everyone else. We must seek out ways of controlling aberrant activities without impeding communication. (13:66)

Networks pose a unique challenge for security and integrity considerations. We must ensure that only the intended destination is the recipient of information transmitted from any point in the network and nowhere else. We must ensure that the information received at any point in a network is the same in content as the data transmitted (nothing added, nothing removed, and not in changed). We must ensure that all components of the network

(terminals, terminal controllers, modems, nodes, data links, and telecommunication lines) on the organizations's premises are accessible only to employees with authorized access. We must ensure that the sender of the information can verify that receipt was by (and only by) the authorized recipient. We must ensure that the recipient of information can verify that the person from whom the communication appears to come is really the person who sent it. We must ensure that information, while in transit, cannot be observed, tampered with, or extracted from the network by some unauthorized person or device. We must ensure that any attempt to observe, tamper, or extract information from the network by an unauthorized person or device can be identified. We do this so that appropriate action can be taken to prevent future occurrences. We must ensure that adequate alternate paths are available to transmit information from any point in a network to any other point to which the need exists. And finally, we must ensure that in the event that a failure of both the primary and alternate communication paths should occur an alternate means of communicating critical information has been identified, implemented, and tested. (14:200-102)

We need to be able to move data and software around a network to the most logical place in the user environment. This enhances productivity and gives the user more flexibility in his capabilities. Distributed database systems can keep track of all information on all databases on all computers in a network, but

they can not guarantee the integrity of the data. "Easy access to data from multiple, heterogeneous remote database management systems could become a reality in the not-to-distant future if an important proposed software standard movement continues to gain momentum." (15:59) Jeff Moad is referring specifically to Remote Data Access (RDA) as a standard protocol for accessing remote databases. The key is that RDA assumes the use of a single, common Structured Query Language (SQL) implementation. There is movement in this direction as "users are beginning to understand the importance of a standard like RDA." (15:63) SQL standardization is not solving all the problems. Eventually, any user of any type of machine in a network will have easy access to data stored on any other machine in that network, no matter which company made the machine or which operating system it uses. RDA could provide a standard way to access databases remotely over a network.

Personal computers have more compute power today than ever before. The trend is to build smaller ones with even more power than some mainframes in current use. Networking these together to only gain access to a central storage facility is not distributed processing or distributed databasing. Although we realize efficiencies in reduction of access time, we do not attain the true potential of the distributed environment. There is a transition as personal computer use expands beyond the personal application and shares software as well as data with the mainframe. Yet, this transition increases cost and compounds the integrity issue.

This whole problem becomes more complicated in the networked multilevel sense in that users at different locations with different authorizations could well be modifying attributes of the same data simultaneously. Networking has only complicated the integrity issue.

What we are discussing is network management. In practice, network management means evaluating hardware technologies with as much emphasis on telecommunications capabilities as on sheer processing performance. It means developing systems level software tools that guarantee network security and create consistent, easy to use interfaces between workstations of different power built around different architectures. It means implementing connectivity standards throughout the organization so that users are free to revise their applications without jeopardizing the company's entire network. Most importantly, the centralized support center must set the technological and organizational ground rules to guide the individual departments and have the authority to enforce those guidelines. In short, it means setting technological and organizational ground rules to guide self-directed computer users.

# CHAPTER VII

## THE DISTRIBUTED ENVIRONMENT

Managers today, are eager to bring the latest technologies into their environments. With the thousands of terminals, workstations, minicomputers, and mainframes now in use, the need to maximize the employment of these resources and maintain integrity of information is compelling. Many believe that distributed data processing means the spread of computer hardware and data to multiple sites within an organization. Distributed processing is really more than what this implies. "The term distributed . . . is properly used to describe a system in which processing is shared among several (or many) workstations, rather than centralized at one location." (16:79) This gets closer to a good definition of distributed processing, but I believe Frederic Withington has captured the true essence of the term. "Real distributed data processing requires the geographical division of a data processing application among multiple sites. It implies intercommunication among the sites for inquires and file updates, and sharing of processing resources, files, and complex data bases." (17:105)

In terms of integrity and security concerns the distributed environment presents serious problems in concurrency control and database modification. Henning and Walker have stated:

> Locking in the distributed environment has to be done very carefully to avoid denial of service to nonlocal nodes which may be doing retrievals against a database while another user is doing updates. . . . The possibility of compromise increases when data is accessed over a distributed system, simply because the user now has access to more than one computer system available for penetration attempts. Denial of service attacks are harder to detect and differentiate from a normal database lock on another node or the time spent in network traffic. The preservation of label integrity and label recognition must also be addressed. It is also possible that the problems associated with data inference and aggregation will become increasingly more complex as additional nodes are added to a distributed system. In addition to all of these problems, the issues of network security must be considered in the development of the distributed database management system. (3:254)

Only now are we beginning to understand the complexities and implications of the distributed database.

> In a system running a distributed data base, not only are there multiple CPU's, but the data as well may be distributed over several mass storage devices located at physically separate sites. The actual location of any item of data does not need to be known in order to make an inquiry, and the process of finding, retrieving, and storing records from the correct mass storage device is completely transparent to the users. (16:79)

The key to successful implementation of a distributed database is the sophistication of the database management executive software and the operating system. Information stored at two or more different sites needs to be treated as a single logical database. The system should be able to resolve the problems that are associated with multiuser databases.

> It must not permit two users to update the same data (the same field or record) at the same time or to carry out conflicting global changes in a file. In addition, the data base software must be able to retrieve the requested data from any physical storage site, update it from any other location, and then transmit the changes back to the point of origin. (16:80)

Frederic Withington's definition acknowledges that data processing is an organizational resource consisting of many areas of activity. Each activity may be executed or controlled by various individuals at various locations within the organization. The act of spreading activities, or areas of responsibilities across an organization is decentralization of computer processing. Managers need to be careful as they try to find out what is appropriate in terms of degree of decentralization for their organizations. The question, how much decentralization, must be resolved to maximize the efficiency and use of the computer resources in a distributed environment. This one issue alone, has created more divisiveness then any other. As such, a closer examination of the causes needs to be undertaken.

We must be cautious as we implement decentralization, for many perils await the unwary. If you allow two or more departments within your organization to develop information systems and write applications, when you want to consolidate reports it becomes difficult if not impossible. Responsiveness to individual needs interferes with corporate level data collection and analysis. To allow individual departments to purchase computer hardware and specify the most needed applications also creates problems. Even if your centralized applications center writes the programs to ensure conformance to standards, backlogs inevitably occur. Frustrated users begin developing applications and buying software products that meet their particular needs. One invariably ends with many

databases and files that are incompatible or cannot be passed to the central computer complex, thus continuity of information and data integrity are lost.

Many organizations still operate in an environment where the centralized development center governs transaction oriented systems and users have limited technical expertise. Applications tend to be simplified yet specialized. The users often lack the clout within the organization to voice or act on any dissatisfaction they feel about the system. The simplicity of the applications and the lack of expertise among end users allows the centralized department to maintain a tight grip on computing systems. It trains and supports users and distributes single, centrally developed versions of software. To minimize response time should be the controlling element not cost when deciding how to meet users' needs.

This environment suffers from several built-in problems. First, the user may spend the bulk of their day working with the computers but they have little control over how they operate. Companies that base service on minimizing costs often create long backlogs for users who seek to have programs updated or modified. Even when cost is not an issue as in many government agencies, there is a resistance for upgrading responsiveness to users. The fear is that too many versions of an application may undermine software consistency. Backlogs and lack of responsiveness breed resentment. Ultimately, the classic information systems dilemma must be faced. Should the centralized department expand its

programming staff to make it more responsive to users needs? Or should it accept the user' growing confidence by loosening control over applications while requiring conformity to data and communications standards. Once users reach a critical mass of restlessness, the status quo is almost impossible to maintain.

Developers need to work closely with users when writing programs and discourage features that might interfere with the company's broad computing goals. Central computing centers can act as software librarians, maintaining programs and swapping applications between offices and departments. Central departments also must play a mediating role as offices and divisions vie for limited resources. Scarce programming resources, particularly for maintaining existing applications, are a great source of instability.

Although a centralized facility can define development procedures in principle, in practice frustrated users threaten program consistency in several ways. Users hire consultants with or without authorization and write programs or entire applications that do not follow guidelines. The availability of low cost, off the shelf software also causes inconsistencies. Users create databases and files that do not meet standards, thus becoming incompatible with existing programs or network software. These pockets of valuable information can introduce inaccuracies into corporate wide data and jeopardize the smooth functioning of strategic applications.

Organizational subunits will differ in goals, time perspectives, interpersonal relationships, and structure. Uncertainty in setting priorities has motivated user group managers to seek control over all system services they see as critical to their operations. When it appears practical and economic, managers who feel sufficiently competent will have strong motivation to control and even to run their own data processing groups.

Power to make decisions often rests at the level where information accumulates and analysis occurs. Since information support is a necessary condition for effective power, managers can use distributed information systems strategically to bolster the authority of system users in the organization.

The specialization of computer applications, causes many organizations to over look potential, and more general, roles for information systems. Information systems are not simply labor saving devices that support the activities of people in one or more departments. They are control and coordination devices that should fit an organization's formal structure and simplify achievement of its goals.

Careful attention must go into planning the arrangement of the data processing resources that develop and operate information systems. Control of activities must be applicable in either developmental or operational environments. Accessing data as a developmental activity usually represents managements' desires. However, they place restrictions on the kinds of data that will be collected and used within the system.

Management can control and coordinate activities not only by direct supervision but also by establishing comprehensive guidelines or standard operating procedures. The control of the technical concerns of database administrators is a typical example. Their responsibilities can be either centralized in an individual or a group. They can be decentralized but constrained by centrally designed standards, or they can be decentralized with virtual independence. In the future, the second approach probably will become increasingly more important. Therefore, the data processing manager can define comprehensive standards that can be enforced by top managers or a highly placed steering committee. These standards can be used in a decentralized organization to protect the data processing department from excessive control by the user.

Central processing units are becoming cheaper as they become more powerful. Although these may be somewhat inexpensive, to maintain redundant peripherals that may sit idle a good bit of the time is not cost effective. With distributed processing, communications costs also may increase as all nodes must be completely interactive.

As we are already aware, integrity errors can occur from undetected erroneous data entry. Errors also occur from software bugs, from equipment or line failures, and deliberate acts. The use of a centralized system simplifies tracing the cause of the error. Unfortunately, the discrepancy is not often caught when it occurs. We need to be prepared to trace the error and restore the system to its proper state.

The need to establish standards has been discussed, but needs to be reemphasized. Distributed processing demands adherence to standards as multiple systems interact in the course of their activities. We must agree upon the definition for data elements. A way to maintain data, whether adding, changing or deleting data is necessary. And, we need to ensure redundancy or recovery of data files if altered or lost. Our most crucial concern is that "consistency on any distributed system is critical." (18:63)

The distributed environment is here to stay and supports the idea of decentralization. Clearly, there needs to be a central authority to ensure that the distributed environment can function and maintain integrity. This becomes critical for the systems employed and the data to be shared.

# CHAPTER VIII

## SUMMARY AND CONCLUSIONS

Information is our most precious resource. As such, the perceptions of the initiator and the recipient decide the value of data and information. When we consider the value of information, we find that it has truly unique characteristics. The same information may have a different value to different people simultaneously and even a different value over time. For example, the same bit of information may be perishable yet timeless. Consider the date and time for the invasion of Normandy. To the planners of the invasion keeping this information secure represented the possible success or failure of the mission. To a German soldier lying in a foxhole on the beaches of Normandy, knowing the date and time of possible invasion could mean life or death. To a historian the date only represents a point of passage or turning point for humanity.

Integrity in a computer system deals with the consistency, accuracy and reliability of information and our ability to create an environment to manage it. Data integrity is concerned with the lowest element of the information chain. Yet, we must be aware that it is possible to have data integrity and not information integrity. Data elements comprise information. By adding to, changing, or deleting the data element, we alter the meaning of information.

Traditional security considerations provide a starting point for the discussion of integrity issues. We need to have a general appreciation for information security. We need to recognize that threats to informational integrity are more likely to occur from accidental or unintentional events. Yet, keep in mind the potential for unauthorized access, modification and destruction of data.

> You can take every precaution and still suffer unauthorized distribution of data, misuse of information, accidental dissemination, or malicious destruction. No stand alone computer systems, local area networks, minicomputers, or mainframes are ever completely safe. All you can do is reduce the chances and minimize the damage. (12:141)

We must begin proper security controls. They include controls in the physical realm, the buildings, the rooms, and the terminals. We need to have access controls, not only to the system but the data that can be obtained through the system. These access controls need to be supplemented with authentication procedures. We must have database and database management system protection. We must protect the connectivity of our workstation, the networks, and communications that support our systems. And, we must institute the management controls to ensure the integrity of the information environment.

Two critical points come to mind. One, the information network will automatically contain errors if standards and levels of precision do not exist. Secondly, without the authority to enforce the standards developed, there is no need for the standard.

The distributed environment is where computer use is moving. This environment encompasses processing and databases which calls for the decentralization of our computer support organizations. We need to be deliberate in our movement away from centralized ideas and ensure that the proper controls are in place to guarantee integrity of data and information.

We must look at the ways we store and retrieve data. We also must be sure to have the right mix of discretionary and mandatory access controls in place. Correctness of data is the primary concern of data integrity in the database management sense. Yet, in the distributed environment, concurrency and access control are serious problems. In order to provide the real-time processing and access we desire, we must solve these problems.

Networks provide us the paths to data and processing that were unattainable just a scant few years ago. Numerous problems in the security and integrity arenas exist because of the openness and accessibility that networks provide. New areas of vulnerability threaten the sanctity of our data, specifically, those points that interface with the terminal, the network, and the communications environment.

We must recognize that information is an asset, as valuable, and as well worth protecting as any other kind of property. We must assess the threat to this asset. We need to determine what kinds of information are vulnerable, to what kinds of threats, and from whom. Finally, we must choose the right techniques and technology to meet the specific threat and challenge of data and information integrity.

# BIBLIOGRAPHY

## Articles and Periodicals


Aaland, Mikkel. "Preventing computer Disaster," Working Women, November 1988, pp. 88(4).

Badgett, Tom. "Data Base Management: Toward a Shared Database." Personal computing, October 1987, pp. 111(4).

Bryan, Marvin. "With High Speed, New Applications," Datamation, December 1, 1988, pp. 59(4).

Carlson, Richard W. "When Words Collide," Vital Speeches, July 15, 1988, pp. 578(6).

Denning, Peter J. "Security of Data in Networks," American Scientist, January-February 1987, pp. 12(3).

Diebold, John. "The Changing Information Environment," Vital Speeches, December 15, 1988, pp. 138(8).

Elmer-Dewitt, Philip. "Don't Tread on My Data," Time, July 6, 1987, p. 84.

Farhoomand, Ali F., and Murphy, Michael. "Managing Computer Security," Datamation, January 1, 1989, pp. 67(2).

Fersko-Weiss, Henry. "Who Manages the Network," Personal Computing, March 1987, pp. 107(6).

Finkelstein, Richard., and Pascal, Fabian. "SQL Database Management Systems," Byte, January 1988, p. 46.

"Government Data Bases and Privacy," Futurist, September-October 1986, pp. 52(2).

Katz, James E. "Telecommunicatins and Computers: Whither Privacy Policy?," Society, November-December 1987, pp. 81(5).

Krause, Bill. "OS-2 and Networking in the 1990's," Personal Computing, October 1988, p. 176.

Lockwood, Russ. "Hypertext Runs Under MS-DOS." Personal Computing, January 1988, pp. 235(2).

McIver, Mary. "Tapping New Secrets," Maclean's, September 28, 1987, pp. 60(2).

Manuel, Tom. "Here Comes Transparent Distributed Computing," Electronics, August 20, 1987, pp. 39(2).

Moss, Christopher D.S. "Intelligent Databases," Byte, January 1987, pp. 97(6).

Nelson, Theodor H. "The Tyranny of the File," Datamation, December 15, 1986, pp. 83(3).

"New Fiber Optic Use," High technology Business, December 1986, pp. 39(2).

O'Malley, Christopher. "Maximizing the Power in Data Base Programs," Personal Computing, November 1987, pp. 147(6).

"Open Systems Interconnection," Fortune, March 27, 1989, pp. 111(12).

Peterzell, Jay. "Spying and Sabotage by Computer," Time, March 20, 1989, pp. 25(2).

Rochell, Carlton C. "The Next Decade: Distributed Access to Information," Library Journal, February 1, 1987, pp. 42(7).

Schultz, Brad. "The Demand for data PBXs Cools as Users,' Datamation, May 1, 1988, pp. 26(2).

Sewell, Alan. "Departmental Computing," Datamation, October 15, 1987, pp. 82(4).

Shulman, Seth. "Greater Awaremess of Security in Aftermath of Computer Worm," Nature, November 24, 1988, p. 30.

_____. "Virus-Proof Computer Security System," Nature, January 5, 1989, p. 4.

Stewart, Ian. "Highly Distributed Processing," Nature, January 5, 1989, p.13.

Verity, John W. "Taming the DASD Monster," Datamation. December 1, 1986, pp. 77(3).

Weissman, Carol. "Global communications and computer Strategies for the 90's," Forbes, September 19, 1988. pp. 61(9).

Books

Chu, Wesley W. Distributed Systems. Dedham, Ma.: Artech House, 1986.

Cooper, James Arlin. Computer and Commuications Security. New York: InterText Publications, 1989.

Daler, T. Security of Information and Data. Chichester, England: Halsted Press, 1989.

Official Documents

National Computer Security Center, Trusted Network Interpretation or the Trusted Computer System Evaluation Criteria, NCSC-TG-005 Version-1, 1987.

U.S. Congress, Office of Technology Assessment. Defending Secrets, sharing Data: New Locks and Keys for Electronic Information. Washington, DC: U.S. Government Printing Office, 1987.

Unpublished Materials

Eggers, Kenneth W., ed. "Trusted DBMS Interptetation of Evaluation Criteria from DoD 5200.28-STD." Fort Meade, Md.: 1988.

Other Sources

Sud, Jagan. Department of Defense, Fort Meade, Maryland. Interview. December 21-27, 1989.

## LIST OF REFERENCES


1. Jackson, Daryl C., Capt. USAF, Air War College, Maxwell AFB, Alabama. Interview, March 15, 1990.

2. Clark, David D. "Evolution of a Model for Computer Integrity." Proceedings 11th National Computer Security Conference. October 17-20, 1988, pp. 14(13).

3. Turn, Rein., ed. Advances in Computer System Security, Volume III, Norwood, Ma.: Artech House, 1988.

4. Henning, Ronda R., and Walker, Swen A. "Data Integrity vs. Data Security: A Workable Compromise." Proceedings 10th National Computer Security Conference, September 21-24, 1987, pp. 334(6).

5. Roberts, Ralph. Compute's Computer security. Greensboro: Compute! Books, 1989.

6. Baker, Richard H. The Computer Security Handbook. Blue Ridge Summit, Pa.: Tab Professional and Reference Books, 1985.

7. Department of Defense Computer Security Center. Department of Defense Trusted Computer system Evaluation Criteria, DOD 5200.28-STD, 1985.

8. Manuel, Tom. "The Assault on Data Security is Getting a Lot of Attention," Electronics, November 1988, pp. 136(4).

9. Peterson, Ivars. "The Complexity of Computer Security." Science News, September 24, 1988, p. 199.

10. Wood, Charles Cresson. "Extended User Authentication: The Next Major Enhancement to Access Control Packages," Data Processing & Communications Security, Spring 1989, pp. 17(5).

11. Schroeder, Matthew T. "What's Wrong with DEMOS," Datamation. December 15, 1986, pp. 66(4).

12. Stranlo, Christine. "The Well Protected Network. Personal Computing, January 1988, pp. 135(6).

13. Morris, James H. "Our Global City," Communications of the ACM, June 1989, pp. 661(2).

14. Report on Information Network Security: The IBM Approach. Delran. New Jersey: Datapro Research, McGraw Hill. (1989).

15. Moad, Jeff. "The Database Dimension," Datamation, May 15, 1989, pp. 59(3).

16. Liskin, Miriam. "Distributing the Workload," Personal Computing. December 1987, pp. 79(4).

17. Witnington, Frederic G. "4 Rules of DDP you Can't Break," Datamation, May 15, 1987, pp. 105(3).

18. Kerr, Susan. "IBM's NFS Alternative," Datamation. January 1, 1989, pp. 63(3).

## GLOSSARY

| | |
|---|---|
| DBMS | Database Management System |
| DoD | Department of Defense |
| I/O | Input/Output |
| LAN | Local Area Network |
| PC | Personal Computer |
| RDA | Remote Data Access |
| SQL | Structured Query Language |